DOCUMENT RESUME

ED 185 683 EA 012 620

AUTHOR Farquhar, Robin H.
TITLE How the Teaching Profession Measures Teaching Effectiveness.
PUB DATE 12 Jun 78
NOTE 13p.: Paper presented at the conference on "Teaching in Medical School" (Saskatoon, Saskatchewan, Canada, June 1978).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Elementary Secondary Education: *Evaluation Methods; *Teacher Effectiveness; *Teacher Evaluation

ABSTRACT

This paper discusses teacher evaluation, including why teachers should be evaluated, what should be evaluated, and how it should be evaluated. Major problems in the evaluation process are listed, followed by a more detailed discussion of the why, what, and how. The paper concludes that dissatisfaction exists in the field of education with current approaches to teacher evaluation. The reason is the failure of research to yield a shared operational definition of teacher effectiveness or adequate instrumentation for reliably measuring and assessing teacher performance. (Author/LD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED185683

HOW THE TEACHING PROFESSION MEASURES TEACHING EFFECTIVENESS

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

by

Robin H. Farquhar

Dean, College of Education

University of Saskatchewan

Presented at the

Conference on "Teaching in Medical School"

Saskatoon

June 12, 1978.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. Farquhar

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

EA 012 620

HOW THE TEACHING PROFESSION MEASURES TEACHING EFFECTIVENESS

I have been asked to present "Education's perspective of what is being done and what should be done" in the area of teacher evaluation. This is a difficult task to accomplish, particularly within the time limit of ten minutes, because there is a substantial body of knowledge, research, and practice in this area. I shall try and capture its essence by, after a brief introduction, focussing my comments on three main questions: Why should teaching be evaluated? What should be evaluated? How should it be evaluated? The who, when, and where of teacher evaluation will be dealt with implicitly.

Extensive study has been undertaken in Education in an effort to determine the qualities and characteristics of "good" teachers, the "best" teaching techniques and procedures, the effects of various methods of teaching on students, the predictability of teacher success, and a variety of other problems related to teacher performance and effectiveness. But the views and evidence provided are often contradictory. They range from those which suggest that teaching quality is impossible to measure because of the complexity of the teaching-learning situation, to those which hold that teacher effectiveness can be determined on the basis of a fifteen-minute observation. That discrepancy indicates some of the problems associated with this area of concern.

More specifically, the major problems in the evaluation of teaching include the following:[1]

1. Definition -- Definitions of good teaching, teacher effectiveness, and teacher competence are often unsupported by research evidence and are as numerous as the number of evaluators.
2. Complexity and variability -- Teaching effectiveness is a difficult area of study because it involves a wide range of human phenomena, qualities, and skills and because the results of teaching must be observed in delayed and inferential fashion. Teaching proficiency is multi-dimensional and there are many kinds of effectiveness for different kinds of teachers, programs, and situations. A teacher functions in a highly complicated setting in which his performance is influenced by the interaction of his personal characteristics and various situational variables. Teacher behaviour varies from one time period to the next according to the nature and purpose of the learning activities. Teachers' classes differ and a fair evaluation of a teacher's performance can only be made on the basis of several years' experience. A teacher's style can only be ascertained if numerous evaluations are made. Situational variables such as characteristics of the physical setting, the instructional objectives, the resources available, the materials being used, and the goals of the institution must also be recognized and considered when evaluating teaching. Also, the ability of the student, his home and social environment, his experiences, and other situations he encounters all influence his learning and the ability of a teacher to teach him. In essence, good teaching is a contextual phenomenon.
3. Research failure -- Studies of teaching effectiveness and teacher evaluation have been inconclusive and do not provide clear

guides to action.

4. Subjectivity -- The evaluation of teaching involves value concepts and value judgements, which grow out of one's personal biases, preferences, beliefs, opinions, and attitudes -- all of which vary substantially from one individual to another. Evaluators tend not to be guided by any common definition of good teaching or theory of learning.

5. Anxiety -- Evaluation is threatening to many teachers, and this generally causes a deterioration in teaching performance (especially in those aspects related to pupil rapport, interaction, warmth, and personal understanding); this often means that the performance exhibited by the teacher is not representative of his best efforts.

6. Who should evaluate? -- Many alternatives have been espoused and used -- evaluation by external sources, by supervisors, by principals or deans, by department heads, by peers, by students, by teams, and self-evaluation. The issues involved here include those of objectivity or psychological distance from the situation versus familiarity with the situation, the use to which the evaluation results are to be put (i.e., professional development and guidance versus decision making and sanctions), the confounding relationships that exist between certain evaluators and the evaluatee (e.g., the administrator's authority and control of sanctions, the colleague's friendship or jealousy, and the student's dependence on the teacher), and the qualifications of the evaluator (including his expertise in evaluation, his knowledge of the instructional goals and situational factors pertaining in a given case, and the availability to him of sufficient

opportunity to evaluate properly).

There is, then, much debate in Education about whether or not teaching can be properly evaluated at all. Nevertheless, it is and must be done. So let's examine the why, what, and how of it.

Why?

There are three main kinds of reasons for evaluating teaching:

1. Public reasons -- The press for accountability demands evidence, in terms of pre-specified goals and objectives, that teachers are fulfilling their responsibility to parents and the general public for providing their children with the best possible learning experiences and ensuring each student a quality education.
2. Professional reasons -- Teachers require information about their performance and diagnostically-based assistance so that their teaching can improve and reach its potential.
3. Administrative reasons -- Data about the effectiveness of teaching are needed as a basis for decisions about the retention, promotion, and salaries of teachers.

These different purposes for evaluation lead to two main types:

1) formative evaluation which is conducted for the purpose of promoting the professional growth and development of the teacher; and
2) summative evaluation, which is conducted for the purpose of official reporting on the effectiveness of the teacher which may lead to judgements regarding the acceptability of his performance.

In Education, the major emphasis in evaluating teachers (the vast majority of whom are tenured) is on bringing about improvement in teaching and learning.

What?

The question of what criteria should be used in evaluating teaching effectiveness remains largely unanswered in Education since no common definition of "good" teaching has been accepted. Nevertheless, three main kinds of criteria have been identified and worked on:[2]

1. Presage criteria -- These are the characteristics with which the teacher comes to the job, and they include such personal and professional traits as personality, appearance, training, and intelligence. Among those that have been found to be related to teaching effectiveness are scholarship in college, professional preparation, attitude toward teaching, and age and experience. Those found to have no significant relationship to teaching effectiveness include intelligence, knowledge of subject matter, cultural background, socio-economic status, sex, and marital status.[3] There are other presage criteria, like personality and appearance, that are much more difficult to measure and I am not aware of research that has yielded significant relationships between them and teaching effectiveness.
2. Process criteria -- These are the aspects of teacher behaviour considered to indicate teaching competence. They involve teaching methods, techniques and strategies, student-teacher relationships, and classroom situations. Interaction analysis and other systematic observation techniques have been developed in an effort to identify various facets of teacher behaviour, social organization, classroom climate and control that correlate with pupil growth. However, we do not yet know enough about classroom interaction to establish the predictive validity of most process variables.

3. Product criteria -- These concern the <u>end result</u> of teaching, and relate to measurable aspects of student growth or change in attitude and behaviour which can be attributed to the effect of the teacher. The major difficulty here lies in demonstrating a direct causal relationship between pupil gain and the behaviour of a particular teacher; this problem includes the difficulty of separating that teacher's effect from those of other teachers and other learning experiences to which the pupil is exposed. Moreover, to the extent that a teacher does have an impact on pupil gain, that gain may be delayed and the impact deferred until some indeterminate future time -- and conversely, an immediately measured learning outcome may not endure for long after the teaching occurs. There is also the difficulty of measuring student gain in the non-cognitive domains of learning. One other caution regarding product criteria concerns the danger that their use will result in teachers' "teaching to the test," which may have some undesirable consequences. To help reduce some of these problems, criterion-referenced testing of student achievement is preferable to norm-referenced testing when the results are to be used as product criteria in the evaluation of teaching.

Educators have placed most of their emphasis in the evaluation of teaching on the process criteria, perhaps because behaviours related to them are most easily observable. However, their validity in predicting pupil growth (the ultimate objective of teaching) has not been convincingly demonstrated.

How?

The most typical methods for evaluating teaching in the field of Education are rating scales and systematic observation (using schedules that focus attention on particular aspects of classroom behaviour such as the teacher's ability to ask high-order cognitive questions, to demonstrate enthusiasm, to use direct and indirect questioning techniques, to probe for student responses, to accept student feelings, to give directions, and to use student ideas). However, observations may not be good measures of teaching effectiveness because the results aren't generalizable beyond the situation under observation, because they are subject to observer bias, because they are expensive in time and money, and because atypical behaviour of the teacher is often demonstrated when the observer is present. Nevertheless, some of these problems can be alleviated through rigorous observer training and the use of video-tape recording. The use of rating scales is also subject to bias, set, and irrelevant factors that intrude in the evaluation process. For example, it has been noted that "a rating scale constructed by those whose view centres on transmission of subject matter is likely to focus almost exclusively on how the teacher organizes and presents subject matter. A scale constructed by one who conceives teaching as guiding learning activities is more likely to stress the skill of the teacher in manipulating the learning environment."[4]

In the vast majority of school systems, classroom observation of teaching by principals or supervisors is the standard method of evaluation. In most cases a check list or rating form is used. The number of observations is typically about three per year for probationary teachers, and perhaps one every two years for tenured teachers. The evaluator sometimes informs the teacher ahead of time that he will be observing a class, and occasionally he may meet formally with the teacher about the evaluation before the observation. While observing, he may take notes of what he sees and hears;

he ordinarily has a rating form of some kind for this purpose. Following the observation, he schedules a conference with the teacher during which he goes over the completed rating form. The reliability of this approach, however, is subject to serious doubt. The results of research conducted in Ontario and New York in recent years demonstrate "conclusively that raters do not show consensus in assigning global scores to the same observation. Even those individuals, principals and supervisors, whose job it is to evaluate personnel and who often claim superior knowledge of what 'good' teaching is, cannot agree among themselves as to the abilities of these particular teachers."[5]

Attempts to systematize teacher observation, based on rationalizing the teaching process, have been made. For example, Flanders[6] has developed an interaction analysis method in which the observer counts various activities occurring in a classroom, such as the number of times the teacher speaks compared to the number of times students speak. Interpreting this ratio for evaluation purposes, however, requires agreement on such matters as the value of student participation and the frequency with which it should occur. Also, an Instrument for Observation of Teaching Activities[7] has been developed in an attempt to break teaching into categories of activities which can be observed, measured, and rated. The problem with such efforts is that, in placing heavy emphasis on rating scales and recorded observations, they require skills based on special training that is costly and time-consuming; moreover, "it is possible to become so caught up in the techniques of evaluation and the language of it that one loses sight of the prime objectives which involve the improvement of instruction."[8]

Conclusion

In general, then, there is dissatisfaction in the field of Education with current approaches to the evaluation of teaching. This is due largely to the failure of research so far to yield a shared operational definition of effectiveness in teaching, a set of criteria that directly relate teacher behaviour to student growth, and adequate instrumentation for reliably measuring and assessing teacher performance.

Nevertheless, we are still working on it. For example, Edward Hickcox at the Ontario Institute for Studies in Education has proposed an alternative approach based on the assumption that there is a clear and strong commitment by the organization to the improvement of teaching.[9] This approach consists of three phases:

1. The objective-setting phase, in which the supervisor and the teacher agree in advance on what the intended outcomes of a period of teaching are, what procedures and resources will be used in the teaching, and what methods and criteria will be employed in assessing the effectiveness of the teaching.
2. The teaching phase, using the procedures and resources agreed to earlier; here, the supervisor may help in providing the support and technical assistance needed.
3. The evaluation phase, conducted according to the methods and criteria agreed on in advance; the teacher is accountable for achieving the learning objectives determined at the beginning, and no new criteria are entered during the process.

At the conclusion of this sequence, the supervisor and teacher together make comparisons between the anticipated and actual outcomes and, if there is a discrepancy, they determine whether it results from unrealistic objectives, from an inappropriate evaluation system, from inadequate

instructional procedures and resources, or from unsatisfactory teaching performance. They make the indicated adjustments and then the process begins again, for it is cyclical in nature.

Such an approach would avoid most of the major weaknesses in typical current efforts at teacher evaluation, but it would require a major commitment by the organization to instructional improvement, particularly in the form of time and talent on the part of the supervisor or instructional developer. While this kind of investment in the measurement of teaching effectiveness is not yet prevalent in the field of Education, it does hold some hope for the future.

So we in Education have not solved the problem of measuring teaching effectiveness. But we have seriously examined its mind-boggling complexity. We would be grateful for any solutions the field of Medicine might provide.

References

[1] Drawn from Adolf John Volk, "Teacher Evaluation in an Urban Saskatchewan School District" (unpublished M.Ed. thesis, University of Alberta, 1972), pp. 12-19.

[2] H.E. Mitzel, "Teacher Effectiveness Criteria," Encyclopedia of Educational Research, 3rd ed. (New York: Macmillan, 1960), pp. 1481-6.

[3] Cf. N.A. Fattu, "Effectiveness -- An Elusive Quality," Education Digest, Vol. 27 (January, 1962), pp. 24-26; and N.A. Flanders, "Teacher Effectiveness," Encyclopedia of Educational Research, 4th ed. (New York: Macmillan, 1969), pp. 1423-37.

[4] W.J. Ellena, M. Stevenson, and H.P. Webb, eds., Question: Who's A Good Teacher? (Washington: American Association of School Administrators, 1961), p. 11.

[5] Edward S. Hickcox, "The Shape of Teacher Evaluation: A Survey of Practices in the Capital District of New York" (unpublished paper, Department of Educational Administration, State University of New York at Albany, 1977), p. 4.

[6] N. Flanders, Analyzing Teaching Behavior (Reading, Mass.: Addison-Wesley, 1970).

[7] National IOTA Council, The Role of the Teacher in Society (Tempe, Ariz.: College of Education, Arizona State University, 1970).

[8] Hickcox, op. cit., p. 23.

[9] Ibid., pp. 25-33.